



TITLE:

# Genetic Algorithmsを用いたパターン検索(カオスとその周辺,研究会報告)

AUTHOR(S):

奈良, 重俊; BANZHAF, Wolfgang

---

CITATION:

奈良, 重俊 ...[et al]. Genetic Algorithmsを用いたパターン検索(カオスとその周辺,研究会報告). 物性研究 1991, 56(2): 235-238

ISSUE DATE:

1991-05-20

URL:

<http://hdl.handle.net/2433/94505>

RIGHT:

## Genetic Algorithms を用いたパターン検索

岡山大学・工学部・奈良重俊

三菱電機（株）・中央研究所・Wolfgang BANZHAF

## §1. はじめに

我々は、生体において実現されているような柔軟な（非ノイマン型）情報処理につき、その情報処理系を非線形動力学系とみなす立場に立って両者の間の構造的に見た関係を探りたいという問題意識を持っている。そしてその見方に立つ物理学的知見を得て、その再構成（応用）の可能性を検討している [1]。

考察に具体性を持たせるためには、考える context を限定することが必要であるが、あまりに限定し過ぎると問題をつまらなくしてしまうため、ある程度汎用性を保った限定を行うのが重要である。従ってここでは、分散表現された記憶のバンク（相空間中のアトラクターセット）からの、曖昧な検索情報をもとにした情報検索を考える。これは不良設定問題（解の存在、あるいは解の一意性が保証されていない）となっている。この類の問題は多次元の相空間において、任意の初期条件から出発して与えられた検索情報（座標を直接指定するものではなく、座標間の複数の曖昧な拘束条件の形で与えられるものとする）を満足する相空間中の一点（あるいは複数点）に至る最短のパス（道程）とそのターゲットを見いだす問題となっている。これをノイマン型の処理アルゴリズムで実行させようとするとは殆どの場合、（1）combinatorial explosion, （2）program complexity を引き起こすものとなっていて、柔軟な情報処理を必要とする例の一つである。これを現実的な処理時間と controle complexity で解こうとする試みは柔軟な情報処理や柔軟な制御といった観点から、あるいは complex dynamics の観点から興味ある問題である。そこにおいて本質的に重要な点は二つあって、

1. 埋め込まれたアトラクターの中でターゲットパターンの basin をいかに早く効率的に見いだすか
2. もしターゲットパターンが記憶中にない場合、ターゲットパターンに近いものを作成（情報の発生）できるか

である [2]。

実際に上の機能を実行させる際に、その記憶検索のための情報処理媒体として、相空間中のランダムウォーク、カオスなどが話題となっている。特にカオスについてはより一般的なものを含め話題が豊富である [3]。しかしながらこの研究会では新たな媒体として最近注目を引いている遺伝子アルゴリズム（Genetic Algorithm あるいは Evolutional Strategy）[4]、[5]、[6] と神経回路網を組合せて用いるやり方とそのシミュレーション結果を報告した。

## §2. 神経回路網

取り上げる具体例としては相空間として、1、-1 の二自由度（興奮と非興奮の二状態）を持つ神経細胞モデルの 400 個の相互結合したネットワーク（ $20 \times 20$  のピクセルから成る画像情報）からなる 400 次元超立方体（hyper cube）の頂点（ $2^{400}$  個）の作る相空間を採用する。そしてこの神経回路網に予め 30 個のパターンを記憶（アトラクター）として埋め

込んでおき、その内の一つをターゲットとして検索 (search) することを考える。

神経回路網の時間発展規則としては一般に用いられている

$$s_i = \Theta \left( \sum_{j=1}^{400} T_{ij} s_j - h \right)$$

を採用することとする。ここに  $s_i = \pm 1$ 、 $T_{ij}$  はシナプス結合行列、 $\Theta$  は階段関数、 $h$  はシナプス結合行列である。記憶させたいアトラクターは、埋め込みたい画像としての興奮パターンに対しその自己相関 (400 個の 1 または -1 の要素から成る状態ベクトルの自己との直積) をとってそれを  $400 \times 400$  のシナプス結合行列とし、それを 30 個分、代数和をとることによって形成した。今回は、記憶させたいパターンとして文字、人の顔、数字等、見て意味のあるパターンを選んだため、パターン間の重なり (overlap) が大きく、そのまま直積をとってアトラクターとすると擬アトラクターが増え過ぎて記憶想起機能が著しく低下する。故に、与えたパターン  $v_\alpha$  の adjoint パターンを

$$v_\alpha^\dagger \cdot v_\beta = \delta_{\alpha\beta}, \quad v_\alpha^\dagger = \sum_\gamma a_{\alpha\gamma} v_\gamma, \quad a = o^{-1}, \quad o_{\alpha\beta} = v_\alpha \cdot v_\beta$$

で定義して、adjoint パターンとの直積をとり、シナプス結合行列を

$$T = \sum_{\alpha=1}^{30} v_\alpha \otimes v_\alpha^\dagger$$

のようにとることにした。このようにある種の直交化の操作を施しても  $T_{ij}$  は対称行列であることは変わらず、系の時間発展においてエネルギー関数がリアプノフ関数となっている事実は不変である。埋め込んだアトラクターのこうした出来上がった basin は極めて大きく、各 basin 中には各アトラクターの検索のたね (種) になる様々な "特徴" が分散して存在していると見て良い。

### §3. 遺伝子アルゴリズム (Genetic Algorithms)

我々は、上記で定義した相空間での有効な情報処理検索軌道を発生させるダイナミックスとして遺伝子アルゴリズムを取り上げて考える。

まず一本の gene としては、評価用のための一個の要素と画像情報要素 400 個の計 401 個の要素から成るものを考える。そしてある定まった数の gene の集合を考え、これを gene pool と呼ぶことにする。この gene pool からランダムに二個あるいは一個の gene を取り出し、その gene に対する操作 (operation) として、cross over (あるいは recombination)、ないしは mutation を行い [7][8]、その新しく得られたパターンを神経回路網に与えてアトラクターに収束させ、所属する basin を調べて与えられた検索情報に基づく評価を行い gene pool に対する入れ替えの基準とした。recombination 及び mutation の具体的操作の詳細については紙面数の関係もあり、省略する [9]。

検索情報が曖昧にしか与えられていないという状況設定下においては、ターゲットパターンの 400 個の要素のうち部分的にしか検索情報が与えられていないという場合を取るべきであるが、今回は簡単のためターゲットについての全情報がわかっているという場合、即ち評価として full pattern matching を採用した。これを本来の partial pattern matching に拡張を行うことは容易であり、次回の発表としたい。

## §4. シミュレーション結果とその考察

最初に行ったシミュレーションは、一試行当り pool 内の gene 数として九個、初期条件としてランダムパターン九個から出発して行った。gene 操作は mutation のみの場合と、recombination と交互に行う場合との二者の場合において、おのおの 7000 から 8000 操作程度を上限として行った、また一操作毎に得られた新たなパターンを神経回路網に入力させ収束結果を評価し、gene pool から取り出したものと比べ、良いものがあれば gene pool に対し入れ替えを行い、悪ければこれを terminate する。これを繰り返して行い、与えられた操作数の範囲内で、九個中一個でもターゲットに一致するものが現れればその試行は成功として二番目の試行に移った。全試行数は 1000 個とし、それらに対し、操作回数の上限  $N$  を決めた時、それに対する成功した試行の個数（成功率） $p(N)$ 、及び微分成功確率  $dp(N)/dN$  を求めた。

比較のため、遺伝子操作一回に対応して、400 個の 1 または -1 からなるランダムパターンを発生させ、その連鎖により相空間中でのランダムウォークを定義して、それを検索軌道としたものをランダムサーチとし、同様な量を計算した。

両者を比較すると、

1. 成功率から言って、遺伝子アルゴリズムが約一桁程度ランダムサーチより優っている。
2. recombination の操作は効率をかなり改善するが、その mutation 操作との相対頻度の最適化を行う必要がある。
3. gene pool 中の遺伝子数を 100 個とすると成功率の数倍の改善が得られた。これはこの方法が並列処理に適していることを示す。
4. シミュレーションの過程で、多数の擬アトラクターが見いだされ、その中には、与えられたアトラクターから意味のある合成画像ができている場合のあることがわかった。

## §5. おわりに

遺伝子アルゴリズムは、高次元相空間中における並列処理のための有用な情報処理媒体となりうるが、その有効性については、他の structural noise による方法（例えば simulated annealing）、あるいはカオスのダイナミックスを用いた方法などとの比較の上で議論する必要があり、現段階では依然として open question である。特に、相空間中を巡る際の軌道の構造について、詳しく解明する必要がある。

## 参考文献

- [1] 奈良重俊、Peter Davis :「非線形動力学と情報処理機能」数理科学、311 (1989) 70 頁
- [2] Peter Davis、奈良重俊 :「カオスと情報処理機能」、合原一幸編「カオス」第七章、サイエンス社 (1990)
- [3] 津田一郎 :「カオスの脳観」、サイエンス社 (1990)

- [4] 池上高志、金子邦彦: Homeo chaotic Symbio-Network、基礎物理学研究所研究会、「カオスとその周辺」、1990年12月、この報告集に同時掲載。
- [5] D.E.Goldberg : *Genetic Algorithms in search, optimization and machine learning* (Addison-Wesley Publishing Company, 1989)
- [6] L.Davis,(ed) : *Genetic algorithms and simulated annealing* (Pitman, London, 1989)
- [7] W.Banzhaf : The "Molecular" Travelling Salesman, *Biological Sybernetics*, vol.64 (1990) 7214
- [8] W.Banzhaf: Optimization by Diploid Search Strategies, *Neural and Synergetic Computers*, vol.42, pp155 (Springer Series in Synergetics edited by H.Haken)
- [9] S.Nara and W.Banzhaf : Pattern Search using Genetic Algorithms and Neural Network, preprint